CDF note 9474

# Measurement of $t\bar{t}$ Cross Section in the Lepton Plus Jets Channel Using Neural Networks in 2.8 $\text{fb}^{-1}$of CDF data

The CDF Collaboration
*URL http://www-cdf.fnal.gov*
(Dated: August 4, 2008)

We present a measurement of the top pair production cross section in $p\bar{p}$ collisions at 1.96 TeV, with an integrated luminosity 2.8 $\text{fb}^{-1}$ at the CDF experiment on the Fermilab Tevatron. We use a neural network technique to discriminate between top pair production and background processes in a sample of events with an isolated, energetic lepton, large missing transverse energy and three or more energetic jets. We measure a top pair production cross section of $\sigma_{t\bar{t}} = 6.80 \pm 0.38$(stat) $\pm$ 0.61 (syst) $\pm$ 0.39 (lumi) pb for a top mass of 175 GeV/c$^2$.

*Preliminary Results for Winter 2006 Conferences*

## I. INTRODUCTION

Since the discovery of the top quark [1], experimental attention has turned to the examination of its properties. Within the context of the Standard Model, in $p\bar{p}$ collisions top quarks are produced in pairs through the strong interaction, via $q\bar{q}$ annihilation (85%) and gluon fusion (15%) at $\sqrt{s} = 1.96$ TeV. Recent theoretical calculations constrain the top pair production cross section with an uncertainty of the order of 8% [2–4].

The top quark is expected to decay to a W boson and $b$ quark nearly 100% of the time. The W boson subsequently decays to either a pair of quarks or a lepton-neutrino pair. Measuring the rate of the reaction $p\bar{p} \rightarrow t\bar{t} \rightarrow \ell\bar{\nu}_\ell q\bar{q}' b\bar{b}$, the lepton+jets channel, tests both the production and decay mechanisms of the top quark.

This note describes a measurement of the top pair production cross section in the lepton+jets channel at $\sqrt{s} =$ 1.96 TeV. We develop a neural network technique to maximize the discriminating power from kinematic and topological variables. The sensitivity of the neural network technique is comparable to that for the traditional CDF secondary vertex $b$-tag method [5, 6], which suppresses the dominant background from $W$+jets at a cost of a 45% loss in signal efficiency. This kinematic method then allows us to check the assumptions in the $b$-tag method and test the modeling of signal and background processes with higher statistics. Exploring the top cross-section in many different channels and using many different assumptions is important for looking for signs of new physics as new physics might appear differently in the various channels. An excellent understanding of top pair production and $W$+jets background kinematics is be required for the searches for the Higgs boson and new physics signatures at both the Tevatron and the future LHC.

## II. DATA SAMPLE AND EVENT SELECTION

This analysis is based on a sample of integrated luminosity of 2.8fb$^{-1}$collected with the CDF II detector between March 2002 and April 2008. The CDF detector is described in detail in [7]. The data are collected with an inclusive lepton trigger that requires an electron or muon with $E_T \geq 18$ GeV ($p_T \geq$18 GeV/c for the muon). From this inclusive lepton dataset we select offline events with a reconstructed isolated electron $E_T$(muon $p_T$) greater than 20 GeV, missing $E_T \geq 20$ GeV and at least 3 jets with $E_T \geq 20$ GeV. In order to suppress the multi-jet background, the leading jet must have an $E_T \geq 35$ GeV and require also that $\not{E}_T \geq 35$ GeV.

### A. $t\bar{t}$ Acceptance

The total acceptance is measured in a combination of data and Monte Carlo. The geometric times kinematic acceptance of the basic lepton+jets event selection is measured using the PYTHIA Monte Carlo program [8]. A top mass of 175 GeV/c$^2$ is used for the acceptance determination. The efficiency for identifying the isolated, high $p_T$lepton is scaled to the value measured in the data using the unbiased leg in $Z$-boson decays. The geometric times kinematic acceptance, is estimated to be 0.0347 ± 0.0003 for central electrons, 0.0185 ± 0.0002 for central muons and 0.0086 ± 0.0001 for forward muons, where the error includes statistical and systematic effects. Table I summarizes the observed number of data events and the expected number of $t\bar{t}$ events assuming a cross-section of 6.7 pb.

| Jet multiplicity | $W \rightarrow e\nu$ | $W \rightarrow \mu\nu$ | Total | Expected $t\bar{t}$ |
|---|---|---|---|---|
| 3 | 2296 | 1604 | 3900 | 532 |
| ≥4 | 781 | 586 | 1367 | 661 |
| ≥3 | 3077 | 2190 | 5267 | 1193 |

TABLE I: The observed number of W candidate events and the expected number of $t\bar{t}$ events, assuming a theoretical cross-section of 6.7 pb at a top mass of 175 GeV/c$^2$.

### B. Backgrounds

The dataset selected above, called "lepton+jets", is dominated by QCD production of W bosons with multiple jets. Much theoretical progress has been made recently to improve the description of the $W$+jets process, with leading-order matrix element generators now available to describe the parton hard scattering for processes with a $W$ and up to six additional partons in the final state. We use the ALPGEN [9] matrix element generator, convoluted with the CTEQ5L parton distribution functions [11]. We require parton $p_T \geq 8$ GeV/c, $|\eta| \leq 3.0$ and minimum separation $\Delta R \geq 0.2$ for $u$, $d$, $s$ and $g$ partons. We have verified that the shapes of the kinematic distributions used in our kinematic analysis are not sensitive to these values. We choose a default momentum transfer scale of $Q^2 = M_W^2 + \Sigma_i p_{T,i}^2$ for the parton distribution functions and the evaluation of $\alpha_s$, where $p_{T,i}$ is the transverse momentum of the $i$-th parton. We use the PYTHIA parton shower algorithm to evolve the final state partons to colorless hadrons. For this analysis, we combine the $W+n$ where n=0,1,..4, parton ALPGEN+PYTHIA Monte Carlo samples to obtain the full kinematic distributions. The dominant contributions come from the W+3p and W+4p samples. These samples are used to model all electroweak backgrounds. The previous version of this analysis [12] showed that the kinematic distributions of these samples are very similar to the W+jets samples. We consider a 1% systematic due to this assumption.

The other substantial background in this analysis comes from events without $W$ bosons. These events are typically QCD multi-jet events where one jet has faked a high-$p_T$lepton and mis-measured energies produce apparent $\not\!\!E_T$. We model the kinematics of this background by using those events that pass all of our selection requirements except lepton isolation [13]. We estimate the rate of such events from a fit to the $\not\!\!E_T$ distribution after all cuts but the $\not\!\!E_T$cut are applied. The QCD background will have predominantly low $\not\!\!E_T$, with tails extending into the signal region. Figure 1 shows the $\not\!\!E_T$distributions used for this fit. In this fit the top-cross-section is constrained to 6.7 pb. This fit tells us that we expect approximately 246 QCD events, which corresponds to a QCD fraction of 4.7%. A 50% relative systematic uncertainty is taken on the number of QCD events. In the final fit, the QCD fraction is constrained to the value obtained by this fit and is constrained with an uncertainty of 50%.

## III. CROSS SECTION MEASUREMENT METHOD

A comparison of the observed data events with the expected number of $t\bar{t}$ signal events in Table I indicates the expected signal to background ratio is about 1:4.4 in the $W+\geq$3 sample. At such low signal purities, the sensitivity to top pair production from the observed number of events alone is eradicated by the large uncertainty on the leading-order theoretical prediction for $W$+jets background. Other CDF measurements of the top pair production cross section have used $b$-tagging, with 55% signal efficiency, to improve the signal-to-background ratio to 2:1 and 3:1, in the $W+\geq$3 jets and $W+\geq$4 jets respectively, and also use the more accurate prediction for the fraction of $W$+jets containing heavy flavor.

This analysis instead exploits the discrimination available from kinematic and topological variables to distinguish top pair production from background. Due to the large mass of the top quark, top pair production is associated with central, spherical, energetic events with different kinematics from the predominantly lower energy background processes. We consider separately two background components: electroweak processes modelled by the $W$+jets Monte Carlo, and multi-jet QCD processes obtained from data. To maximize our discriminating power, we use an Artificial Neural Network (ANN) technique [14]. ANN's employ information from several variables while accounting for the correlations among them.

We perform a binned likelihood fit to the discriminating variable and find the most likely number of events from $t\bar{t}$ production, $n_{t\bar{t}}$:

$$L(\sigma_{t\bar{t}}, n_w, n_q) = \Pi_{i=1}^{N_{data}} \frac{e^{n_i} n_i^{d^i}}{d_i!}$$

where $\sigma_{t\bar{t}}$, $n_w$, $n_q$ are the parameters of the fit, representing the $t\bar{t}$cross-section, the number of $W$-like and multi-jet events respectively present in the sample. Moreover, the expected number of events in the $i$-th bin is

$$n_i = (\sigma_{t\bar{t}} \cdot \frac{\epsilon_{t\bar{t}}}{\mathcal{L}} P_{t\bar{t},i} + n_w P_{w,i} + n_q P_{q,i}), \tag{1}$$

where $P_{t\bar{t},i}$, $P_{w,i}$, $P_{q,i}$ are the probability of observing an event in bin $i$ from $t\bar{t}$, $W$-like and multi-jet processes. $\epsilon_{t\bar{t}}$ is the acceptance estimat including the branching ratio for $W \to \ell\nu$, and $\mathcal{L}$ the luminosity measurement. $zi$ denotes the number of observed data events that populate the $i$-th bin. The level of the multi-jet background, $n_q$ is fixed to that expected from the fit to the $\not\!\!E_T$distribution with an uncertainty of 50%.

### A. Neural Network

The Neural Network method used for the previous version of this analysis was maintained. There are many algorithms one could use for adjusting the weights in order to produce an optimized network [15]. For this particular problem, the previous version of this analysis, obtained satisfactory results by using the default JETNET back-propagation training method with a term added to the error function in order to discourage large weights.
The same 7 inputs to the ANN were chosen for this analysis. The variables of choice are shown in table II

| Variable | Definition |
|---|---|
| $H_T$ | Scalar sum of transverse energies of jets, lepton and $\not{E}_T$. |
| Aplanarity | $3/2Q_1$ |
| $\Sigma p_z/\Sigma E_T$ | Ratio of total jet longitudinal momenta to total jet transverse energy. |
| $\min(M_{jj})$ | Minimum di-jet invariant mass |
| $\eta_{max}$ | Maximum $\eta$ of jet. |
| $\Sigma_{i=3}^{5} E_{T,i}$ | Sum $E_T$ of third, fourth and fifth jets. |
| $\min(\Delta R_{jj})$ | Minimum di-jet separation in $\eta$ and $\phi$. |

TABLE II: Definition of variables used as inputs to the ANN this analysis. The momentum tensor of the event is formed from the lepton, $\not{E}_T$ and the $E_T$ of the five highest $E_T$ jets. The eigenvalues are ordered such that $Q_1 \leq Q_2 \leq Q_3$.

The ANN is a feed-forward perceptron with one intermediate (hidden) layer and one output node. For training, we use 5000 ALPGEN $t\bar{t}$ and 5000 ALPGEN+PYTHIA $W$+jets Monte Carlo events and require an output of 1.0 for $t\bar{t}$ signal and 0.0 for $W$+jets background. Other sources of background are not considered during the training process. The weights of the network are adjusted to minimize a typical mean squared error function:

$$E = \frac{1}{N}\sum_{i}^{N}(O_i - t_i)^2$$

where $O_i$ is the output of the network for the input event $i$ and $t_i$ is the desired target value. The target values are 1.0 for signal and 0.0 for background events. Learning is an iterative process and we use an independent testing sample of 1900 PYTHIA $t\bar{t}$ and 1900 ALPGEN+PYTHIA $W$+jets Monte Carlo to evaluate the ANN performance and choose when to stop training. After training was completed, an independent validation sample was used to check the quality of the training.

## IV. SYSTEMATIC UNCERTAINTIES

Systematic uncertainties in this analysis come from Monte Carlo modeling of the geometrical and kinematic acceptance for signal, the luminosity measurement, and from modeling of the kinematic shapes for signal and background. The list of the systematic uncertainties we have considered is summarized in Table III.

1’500’000 pseudo-experiments (PE) are thrown using the shifted templates (both shape and normalisation are changed when relevant). We fit the PE using the nominal templates. Note that for the systematics that affect the W+jets, as the normalisation is left to float in the fit, the only systematic effect considered is that of the shape. The systematics affecting the shape of the W+jets are: the jet energy scale (JES) and interaction scale ($Q^2$). For cases where we have two shifts representing an increase and a decrease of some variable, the systematic is determined as half of the maximal deviation between the three values: nominal, shifted up and shifted down.

The largest source of systematic uncertainty comes from the shape of the QCD background. The the shape of the QCD background is strongly dependent on the particular models used to model it. We use non-isolated leptons to model as the default model. Other models considered are the anti-electron model and the jet-electron model. For the anti-electrons, we require the electron to fail 2 out of 5 selection cuts (not including isolation). The statistics for this samples are too small to use this model as the default but the shapes appear to provide the best modelling. In the jet-electron model we consider a sample of event passing an inclusive di-jet trigger and require the events to pass the same event selection as our sample. The contamination due to $t\bar{t}$is considered to be small due to the heavy pre-scaling of this sample. The observed shift in the measured cross-section is 6.2% for the jet-electron sample and 1.4% for the anti-electron sample. In order to be conservative, we consider the deviation due to the jet-electron sample.

The next largest uncertainty, denoted $t\bar{t}$generator uncertainty in Table III is due to the difference when comparing the $t\bar{t}$signal modelling from HERWIG and from PYTHIA Monte Carlo. The kinematic distributions do not seem to be very different but there is a significant change in acceptance between these two samples.

The jet energy scale uncertaintly comes from the uncertainty on the jet energy corrections for different calorimeter response (as a function of $\eta$), the absolute hadron energy scale, and fragmentation etc. This affects simultaneously five of the seven kinematic variables used in the ANN analysis.

The initial- and final-state radiation uncertainties (IFSR) on $t\bar{t}$are estimated by increasing and decreasing simultaneously the parton shower evolution parameters by an amount based on studies of the CDF Drell-Yan data. The effect of the interaction scale variation, $Q^2$, on the W+jets background is estimated using a new CDF prescription for varying some ALPGEN parameters. The PDF uncertainties on $t\bar{t}$are obtained by considering 46 sets PDF eigenvectors.

The luminosity of 2.8 fb$^{-1}$ has an uncertainty of 5.8%, of which 4.2% comes from the acceptance and operation of the luminosity monitor and 4.0% from the calculation of the total $p\bar{p}$ cross section [16].

| Effect | Acceptance | Shape | Uncertainty W+≥3 jets (%) ) |
|---|---|---|---|
| Jet $E_T$ Scale | YES | YES | 3.1 |
| W+jets $Q^2$ Scale | NO | YES | 2.6 |
| $t\bar{t}$ IFSR | YES | YES | 0.8 |
| QCD shape | NO | YES | 6.2 |
| QCD fraction | NO | YES | 1.4 |
| $t\bar{t}$ generator | YES | YES | 4.6 |
| $t\bar{t}$ PDF | YES | NO | 0.5 |
| Other EWK | NO | YES | 1.0 |
| Lepton ID/trigger | YES | NO | 0.6 |
| Zvtx SF | YES | NO | 0.2 |
| Total before Lumi | - | - | 8.9 |
| Luminosity | YES | NO | 5.8 |
| Total | - | - | 10.6 |

TABLE III: Table of the systematic uncertainties. The overall uncertainty is obtained by adding in quadrature the individual effects.

## V. RESULTS

For $t\bar{t}$ events in 3 or more jets, assuming a top mass of 175 GeV/c$^2$we measure a cross section with the artificial Neural Network technique of

$$\sigma_{t\bar{t}} = 6.80 \pm 0.38(stat.) \pm 0.61(syst.) \pm 0.39(lumi.). \tag{2}$$

where the first uncertainty is statistical, the second is systematic excluding the luminosity and the third is the luminosity uncertainty. These results are in good agreement with the theoretical prediction of 6.7 pb for a top mass of 175 GeV/c$^2$. The expected statistical sensitive was estimated using 1'500'000 pseudo-experiments and was found to be 0.380. The observed statistical sensitivity is compatible with this value. The NN output distribution used for the final fit is shown in Figure 2.

### Acknowledgments

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of

China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucleaire et Physique des Particules/CNRS; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, Spain; the Slovak R&D Agency; and the Academy of Finland.

[1] CDF Collaboration, F.Abe *et al.*, Phys. Rev. Lett. **74** 2626 (1995); D0 Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **74**, 2632 (1995).
[2] M. Cacciari, S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, [arVix:0804.2800] (2008) and JHEP **0404**, 068 (2004) [arXiv:hep-ph/0303085].
[3] N. Kidonakis and R. Vogt, [arXiv:0805.3844] (2008) and Phys. Rev. D 68 114014 (2003).
[4] S. Moch and P. Uwer, [arXiv:0807.2794] (2008).
[5] CDF collaboration, "Measurement of the Top Pair Cross Section in the Lepton plus Jets decay channel with 2.7 $fb^{-1}$", CDF public note 9462
[6] CDF collaboration, "Measurement of the $t\bar{t}$ productin cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using Lepton+Jets Events with Secondary Vertex b-tagging", Winter 2006 public conference note.
[7] F. Abe, *et al.*, Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988); D. Amidei, *et al.*, Nucl. Instum. Methods Phys. Res. A **350**, 73 (1994); F. Abe, *et al.*, Phys. Rev. D **52**, 4784 (1995); P. Azzi, *et al.*, Nucl. Instrum. Methods Phys. Res. A **360**, 137 (1995); The CDFII Detector Technical Design Report, Fermilab-Pub-96/390-E
[8] T. Sjostrand *et al.*, High-Energy-Physics Event Generation with PYTHIA 6.1, Comput. Phys. Commun. **135**, 238 (2001).
[9] M.L. Mangano *et al.*, JHEP **07**, 1 (2003).
[10] G. Corcella *et al.*, HERWIG 6: An Event Generator for Hadron Emission Reactions with Interfering Gluons (including supersymmetric processes), JHEP **01**, 10 (2001).
[11] H.L. Lai *et al.*, Eur. Phys. J. **C12**, 375 (200).
[12] CDF collaboration, "Measurement of the Cross Section for $t\bar{t}$ Production Using Event Kinematics in $p\bar{p}$ Collisions at 1.96 TeV", CDF public note 8092.
[13] Lepton isolation is defined as the excess transverse energy in the calorimeter within a cone of radius 0.4 centred on the lepton direction, divided by the lepton transverse energy. The selection requirement is less than 10%.
[14] Brian D. Ripley, *Pattern recognition and Neural Networks*, Cambridge University Press (1996).
[15] Cartsen Peterson, Thorsteinn Röngnvaldsson, Leif Lönnblad, *JETNET 3.0 - A Versatile Artificial Neural Network Package*, CERN-TH.7135/95.
[16] S. Klimenko *et al.*, FERMILAB-FN-0741 (2003); D. Acosta *et al.*, Nucl. Instrum. Meth. **A494**, 57 (2002).

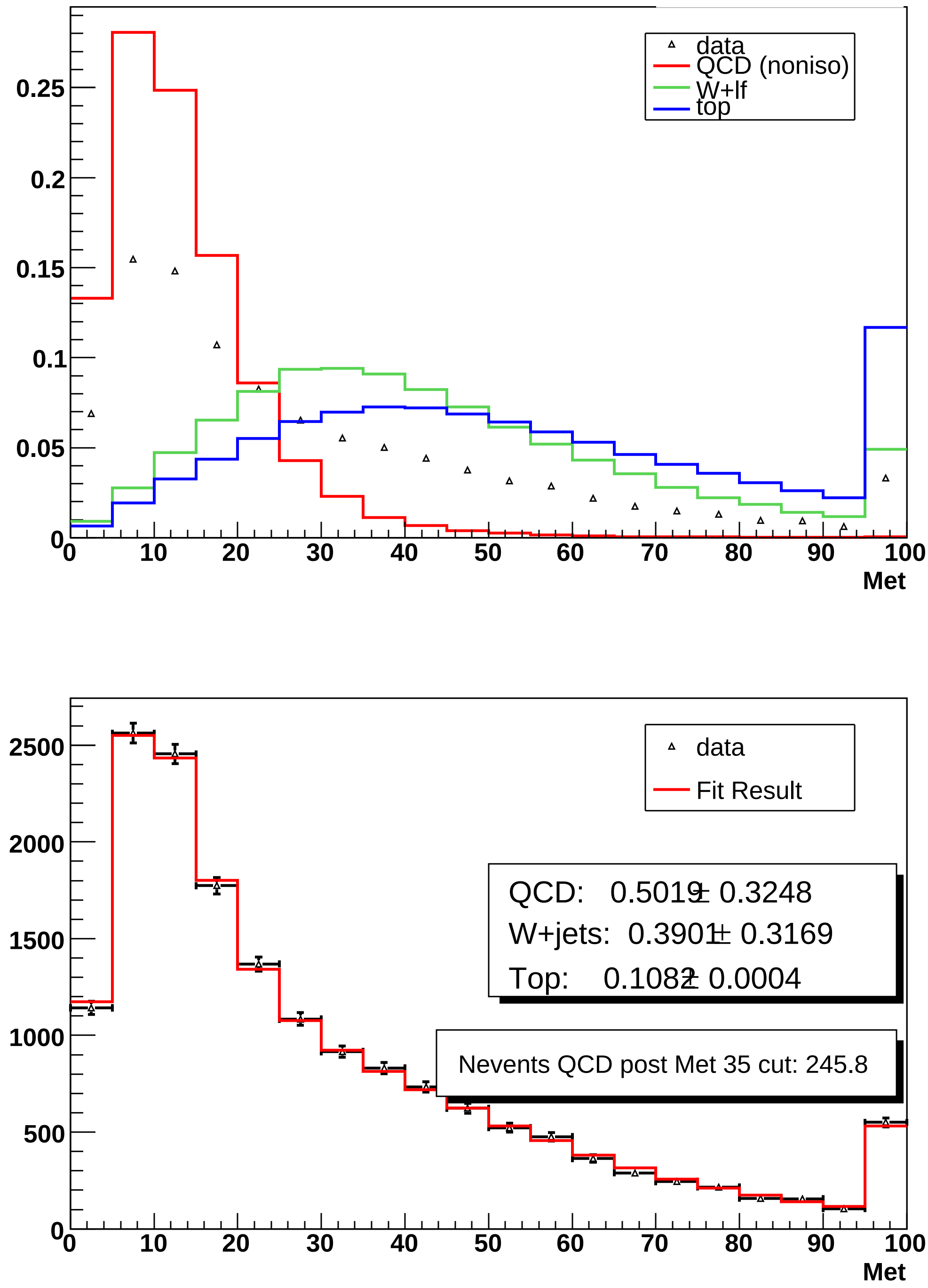


FIG. 1: (top) $\not\!\!E_T$templates for the data, top signal, W+jets and QCD backgrounds in the W+≥3 jet case. These plots are normalised to unit area. (bottom) Comparison between the data and the fitted distribution. The fractions shown in the legend are before any $\not\!\!E_T$cut is applied.

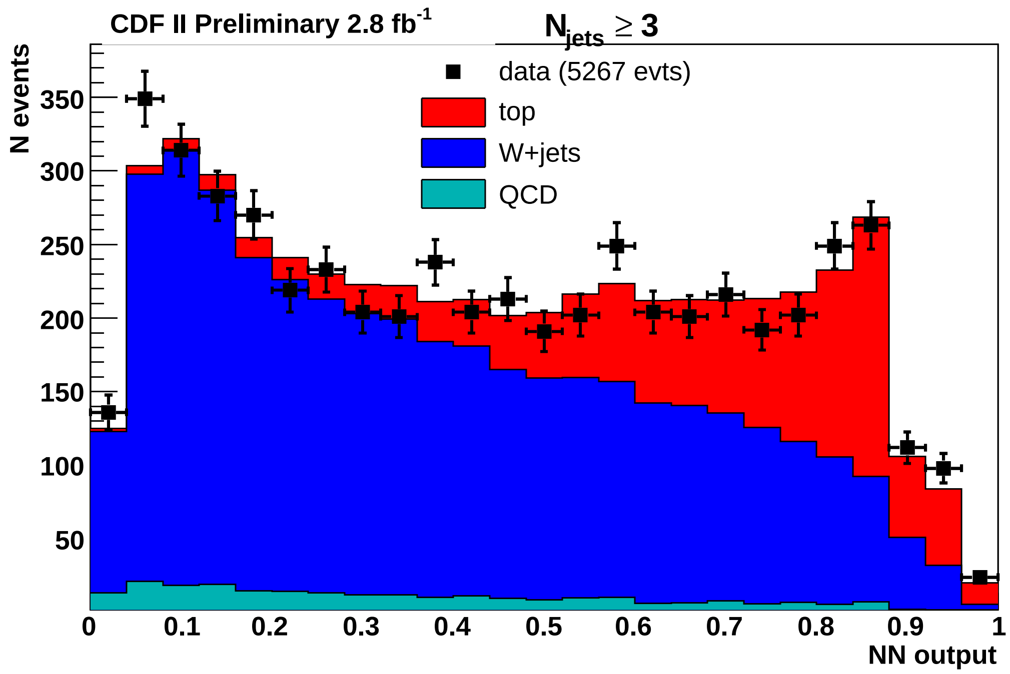


FIG. 2: Fit to the NN output variable in the $\geq 3$ jet sample.